Investigation of Synbiotic Treatment in NAFLD (INSYTE)

Statistical Analysis Plan

Date: 4th June 2014

Version: 1.0

Prepared by	**Signature**	**Date**
Helen Moyses RBRU/NBRC Statistician	_____	_____

Approved by:	**Signature**	**Date**
Christopher Byrne Principle Investigator	_____	_____
Eleanora Scorletti Clinical Research Fellow	_____	_____

Document history

Version	Date	Changes made

1.0 Outline

Trial design: This is a randomised, single centre, double-blind clinical trial.

Trial objectives: The trial addresses the hypothesis that treatment with a synbiotic will have a beneficial effect in non-alcoholic fatty liver disease (NAFLD) compared to placebo, as measured by change in liver fat and serum biomarkers.

Participants: Patients aged >18 years with NAFLD, either biopsy-proven or confirmed by non-invasive imaging in a high-risk cohort (i.e. diabetic and/or features of metabolic syndrome). Alcohol consumption ≤ 14 units / week for women ≤ 21 units / week for men.

Intervention: A dose of a synbiotic nutritional supplement comprising fructo-oligosaccharide with a degree of polymerization < 10 at 4 g/twice a day (two sticks a day) plus Bifidobacterium animalis subsp. lactis BB-12 at a minimum of 10 billion CFU/day (1 capsule a day) will be taken for 10-14 months.

Treatment comparison**:** Synbiotic versus placebo

Primary outcomes:

i. To assess change in liver fat with synbiotic treatment

ii. To assess change in biomarkers with synbiotic treatment

iii. To assess change in gut microbiota composition with synbiotic treatment

Secondary outcomes:

i. Liver fibrosis determined by transient elastography

ii. Insulin and glucose concentrations and hepatic insulin sensitivity

iii. Microvascular function

iv. Plasma cardiovascular risk markers

v. Adipose tissue and blood markers of metabolism and inflammation

vi. De novo lipogenesis

vii. Satiety and satiety factors

viii. Intestinal permeability

ix. Carotid intima-media thickness

x. Pulse wave velocity

xi. Ankle brachial pressure index

Study site: University Hospital Southampton

2.0 Sample size and power

There is little published literature to date upon which to base a sample size calculation to test the effects of the synbiotic treatment on the primary end points. However, the study has been powered on the basis of the expected change in the key primary end point of a change in liver fat.

A sample size of 50 in each group, with a 14% drop out during the study, will have 85% power to detect a difference of 40% in liver fat (in the treatment arm compared with the placebo), assuming that the common standard deviation is 62%, using a power calculation test with a 0.05 two-sided significance level. In our recent randomised controlled trial over a longer period of intervention, 5% of the randomised cohort withdrew between randomisation and end of study measurements. In this recent trial, the mean percentage fat content was 28.5% with a similar standard deviation to that presented above. Liver fat content will be measured by MRS spectroscopy and this technique can reproducibly detect liver fat content as low as 5%. One recent publication in people with NAFLD testing the effect of a synbiotic over 24 weeks in a small study showed a 69% decrease in liver fat.

3.0 Randomisation procedures

A randomisation list will be generated by the BRU Data Management and Statistics Group. Participants will be stratified by gender, age (<50 and >=50) and whether they agree to participate in the sub-study, and blocks will be used to ensure a balance between treatment groups within strata.

4.0 Data monitoring and interim analyses

Detailed analysis and publication of baseline data will be undertaken prior to trial completion by Profs, Byrne and Calder and other investigators in discussion with the research nurses and RBRU/NBRC statistician. An

independent data monitoring committee (IDMC) has not been set up for this trial, so there are no planned interim analyses of outcome data.

5.0 Procedures for data checking

Data will be checked as described in the INSYTE Data Management Plan. Prior to any statistical analysis, all variables will be checked for number of missing values, impossible and improbable values. Impossible and improbable values for continuous variables will be defined by clinical opinion, other improbable values, defined as being more than three standard deviations from the mean, will also be checked. Errors in categorical variables will be identified as those with values missing from given terms in the Data Dictionary. Inconsistencies will be referred to a named investigator by the Data Manager. Descriptive statistics will be calculated for all variables and distributional assumptions checked.

6.0 Trial analysis

Data will be analysed by Christopher Byrne and Eleanora Scorletti, unless otherwise stated.

6.1 Description of baseline data

All important variables collected at baseline will be summarised by treatment group. For continuous variables; means and standard deviations will be calculated. Distributions will be assessed using histograms and the normality assumption will be tested. When the data are not normally distributed medians and percentiles will be used to summarise data. Number and percentage will be calculated for categorical and binary data.

6.2 Primary endpoint

Definition

Change in liver fat and biomarkers from randomisation to end of study.

1. Change in liver fat (measured by MRS or IPOP MRI).The study is powered for this endpoint.

2. Change in biomarkers (not powered for)

a) ELF score which is calculated using the following algorithm:
Score= -7.412 + (ln(HA)*0.681) + (ln(P3NP)*0.775) +
(ln(TIMP1)*0.494) + 10.

b) NAFLD fibrosis score which is calculated using age, hyperglycemia, body mass index, platelet count, albumin and AST/ALT ratio.

c) CK18 M30 and M65 as markers of apoptosis and necrosis

3. Change in gut microflora assessed by 16S rRNA, FISH and qPCR (not powered for).

Measurement

1. Magnetic resonance spectroscopy (MRS) and/or MRI (IPOP) imaging will be used to determine the amount of liver fat as a percentage of the whole liver. This will be performed at baseline and at the end of study visit.
2. Biomarkers will be measured in plasma or serum or will be generated from an algorithm.
3. Gut microflora will be assessed by 16S rRNA, FISH and qPCR.

Statistical analysis

There are three primary endpoints; change in liver fat, change in biomarkers and change in gut microflora.
For measurements of liver fat, biomarkers and gut microflora:
Multiple regression analysis will be used, with final visit measurement as the outcome, and baseline measurement and randomised treatment group as predictors.
For each measurement, a multiple regression analysis will also be performed adjusted for potential confounders.

<u>Potential confounders include but are not limited to:</u>

Change in body weight

Age

Sex

Physical activity

For all outcomes, the treatment effect with 95% confidence interval will be reported (regardless of whether statistical significance is met). This will allow discussion of whether the trial result is compatible with a clinically important effect.

When reporting results it will be acknowledged that care should be taken in interpreting results from several outcome measures, as some statistically significant findings are likely to result from chance alone; consequently, we will allow for multiplicity of testing when interpreting our results.

Analysis populations and missing data

The primary analysis will be: a) intent-to-treat (ITT) and b) per protocol analysis.

1. The ITT analysis will include all patients with complete data (i.e. having baseline and end of study measurements) in the groups to which they were randomised (regardless of whether they were later found to be ineligible, a protocol violator, given the wrong treatment allocation or never treated). This analysis assumes that any missing data is missing at random (MAR), i.e. that there is no difference between missing and observed values, once adjusted for any baseline variables which predict for missingness.

In order to explore the effect of departures from the MAR assumption, sensitivity analyses will be performed.

This assumes that d= the mean of the missing data minus the mean of the observed data. Under MAR, d=0. The value of d will then be varied in order to model different scenarios (i.e. that the patients who are lost to follow up have systematically worse outcomes), and we will report whether the significance of the main analysis is maintained in the sensitivity analysis.

All randomised participants will be accounted for in these analyses.

2. Per protocol analysis will include all patients who consumed ≥50% of their supplement in the time period from randomization to final visit, and will exclude participants who were later found to be ineligible or who did not complete the study.

3. Investigations are currently underway to identify a marker of change in gut microbiota. Once this has been identified, a secondary analysis will be performed using this marker as the explanatory variable (rather than randomisation group) and liver fat (or biomarkers) as outcome variables. This will be elaborated on in later versions of this document.

6.3 Secondary endpoints

6.3.1 Liver fibrosis determined by transient elastography

Definition

This is a new non-invasive, painless method allowing the evaluation of liver fibrosis. The probe utilised for the elastography is an ultrasound transducer mounted at the end of a vibrating cylinder. The cylinder produces a vibration that is transmitted towards the tissue. The ultrasound detects the propagation of the vibration by measuring its velocity. The velocity of the wave towards the tissue is directly related to the tissue stiffness. High velocity is related to higher tissue stiffness that corresponds to an increased severity of fibrosis. The performance of the transient elastography has been assessed in a meta-analysis including fifty studies.

Measurement

We will undertake transient elastography at the beginning and at the end of the study to measure liver stiffness.

Statistical analysis

Each patient's change from baseline to end of study visit will be calculated. If normally distributed, the mean difference (in change from baseline) between the two groups will be calculated with 95% confidence interval. If not normally distributed, log transformation will be used if appropriate. Alternatively, the median, interquartile range and minimum/maximum values will be reported.

If appropriate, a regression analysis will be performed including baseline and treatment group as explanatory variables, in order to estimate the means for both groups (adjusted for baseline).

6.3.2 Insulin and glucose concentrations and hepatic insulin sensitivity

Measurement

The hyperinsulinaemic euglycaemic clamp, which is the gold standard for measuring insulin sensitivity, will be performed on a subset of 24 patients. The quantity of exogenous glucose required to maintain euglycaemia (the M-value) will be calculated at baseline and at end of study.

Statistical analysis

Each patient's change from baseline to end of study visit will be calculated. If normally distributed, the mean difference (in change from baseline) between the two groups will be calculated with 95% confidence interval. If not normally distributed, log transformation will be used if appropriate. Alternatively, the median, interquartile range and minimum/maximum values will be reported.

If appropriate, a regression analysis will be performed including baseline and treatment group as explanatory variables, in order to estimate the means for both groups (adjusted for baseline).

6.3.3 Microvascular function - Laser Doppler Fluximetry

Definition

Two small laser Doppler probes (Class 1 2.5mW maximum) will be positioned, one on the volar surface of the non-dominant forearm and one on the index finger, to detect blood flow in the superficial dermalvasculature. A blood pressure cuff will be placed around the upper arm and blood flow measured before, during and after inflation of the cuff to suprasystolic pressure maintained for up to 3 minutes. The reactive hyperaemic response will be used to assess the capacity of the vasculature to dilate under rested conditions. Spectral power density of the Doppler flux signals will be analysed off line using

existing software packages to determine the relative power density attributed to vasomotor activity within the microvascular bed.

Measurement

Measures to be recorded include the following where flux is an arbitrary unit of blood flow in the dermalvasculature:
Deep Inspiratory Gasps DIG (alteration in flux after DIG, mean of 3)
Resting Flux RF (average flux prior to inflation)
Maximum Flux MF (maximum flux after release of cuff pressure)
MF/RF Ratio (maximum flux over resting flux)
Area of Hyperaemia AH (area between resting flux and flux trace)
Time to Maximum Flux TM (interval between release and maximal flux)

Statistical analysis

Each patient's change from baseline to end of study visit will be calculated. If normally distributed, the mean difference (in change from baseline) between the two groups will be calculated with 95% confidence interval. If not normally distributed, log transformation will be used if appropriate. Alternatively, the median, interquartile range and minimum/maximum values will be reported.
If appropriate, a regression analysis will be performed including baseline and treatment group as explanatory variables, in order to estimate the means for both groups (adjusted for baseline).

6.3.4 Plasma cardiovascular risk markers

Definition

A number of plasma cardiovascular risk and prognostic markers will be analysed.
Examples include:

1. NT-ProBNP
2. hsCRP
3. MMPs

4. Endothelial microparticles
5. TIMP-1
6. ICAM
7. VCAM
8. Interleukins and other cytokines

Measurement

These plasma biomarkers will be measured at the beginning and end of study.

Statistical analysis

Each patient's change from baseline to end of study visit will be calculated. If normally distributed, the mean difference (in change from baseline) between the two groups will be calculated with 95% confidence interval. If not normally distributed, log transformation will be used if appropriate. Alternatively, the median, interquartile range and minimum/maximum values will be reported.
If appropriate, a regression analysis will be performed including baseline and treatment group as explanatory variables, in order to estimate the means for both groups (adjusted for baseline).

6.3.5 Adipose tissue and blood markers of metabolism and inflammation

Definition

G protein-coupled receptor 43 (GPR-43) is a protein expressed in the gut and adipose tissue and is implicated in lipolysis regulation and adipocyte differentiation. A high fat diet increases the expression of GPR-43 mRNA in the adipocytes. It has been demonstrated that synbiotic supplementation counteracts this effect and inhibits lipolysis.

Measurement

Insulin sensitivity (NEFA and glycerol)
Lipolysis (NEFA and glycerol)

Macrophages infiltration
Immune cells
Inflammatory gene expression (have to decide which genes)
Inflammatory proteins
Adipose tissue fibrosis
Quantification of blood vessels
Fatty acid composition
Adipocyte size

Statistical analysis

Each patient's change from baseline to end of study visit will be calculated. If normally distributed, the mean difference (in change from baseline) between the two groups will be calculated with 95% confidence interval. If not normally distributed, log transformation will be used if appropriate. Alternatively, the median, interquartile range and minimum/maximum values will be reported.
If appropriate, a regression analysis will be performed including baseline and treatment group as explanatory variables, in order to estimate the means for both groups (adjusted for baseline).

6.3.7 Satiety and satiety factors

Definition

We will measure satiety at the beginning and at the end of the study. Participants will be instructed to abstain from alcohol and strenuous physical activity for 2 days prior to the day of the test. We will offer participants a free-choice buffet breakfast (comprising yogurt, bread, butter, cheese, jam, fruit, orange juice, and water; approximately 470 kcal); participants will be instructed to complete breakfast within 15 minutes. Food and drink will be weighed before and after the meal and we will calculate energy intake. Before and after breakfast (at -5, 0, 15, 30, 60, 120, 180 minutes) we will assess appetite rating, using a 100-mm visual analogue scale; moreover, we will measure gut hormones (blood sample).

Measurement

We will use a satiety visual analogue scale with verbal descriptors expressing the most positive and the most negative ratings positioned at each end of a 100-mm line. We will ask participants to draw a vertical mark across the line corresponding to their feelings from 0 (not hungry at all) to 100 (very hungry). We will quantify the level of satiety by measuring the distance from the left end of the line to the mark.

We will measure gut hormones (blood sample): glucagon-like peptide 1 (GLP-1), glucose-dependent insulinotropic polypeptide (GIP), peptide YY (PYY), and/or ghrelin.

Statistical analysis

The area under the curve will be used to describe the satiety visual analogue scale over time.

The following measures will be used to describe the gut hormone levels over time:

Gut hormone levels at -5 minutes (fasting) and 180 minutes.

Area under the curve.

Peak gut hormone level

Time to peak gut hormone level.

Each patient's change from baseline to end of study visit will be calculated. If normally distributed, the mean difference (in change from baseline) between the two groups will be calculated with 95% confidence interval. If not normally distributed, log transformation will be used if appropriate. Alternatively, the median, interquartile range and minimum/maximum values will be reported.

If appropriate, a regression analysis will be performed including baseline and treatment group as explanatory variables, in order to estimate the means for both groups (adjusted for baseline).

6.3.8 Intestinal permeability

Definition

After an overnight fast, we will administer to each participant a solution containing 10 g of lactulose and 5 g of mannitol in 35 ml of water (1300 mOsml/L). Urine will be collected over the next 6 hours in plastic containers with 1 ml of chlorexidine, 2% as preservative, to prevent bacterial degradation of sugars.

Measurement

Therefore, outcome is measured by the levels of the 2 sugars in the urine samples collected over the 6 hours.

Statistical analysis

Each patient's change from baseline to end of study visit will be calculated. If normally distributed, the mean difference (in change from baseline) between the two groups will be calculated with 95% confidence interval. If not normally distributed, log transformation will be used if appropriate. Alternatively, the median, interquartile range and minimum/maximum values will be reported.

If appropriate, a regression analysis will be performed including baseline and treatment group as explanatory variables, in order to estimate the means for both groups (adjusted for baseline).

6.3.9 Carotid intima-media thickness with treatment

Measurement

Carotid intima-media thickness (CIMT) is a non-invasive test in which the internal lining (intima-media layer) of the carotid arteries will be measured with B-mode ultrasound.

Statistical analysis

Each patient's change from baseline to end of study visit will be calculated. If normally distributed, the mean difference (in change from baseline) between the two groups will be calculated with 95% confidence interval. If not normally distributed, log transformation will be used if appropriate. Alternatively, the median, interquartile range and minimum/maximum values will be reported.

If appropriate, a regression analysis will be performed including baseline and treatment group as explanatory variables, in order to estimate the means for both groups (adjusted for baseline).

6.3.9 Pulse wave velocity

Measurement

This will be undertaken using radial applanation tonometry and SphygmoCor software to derive non-invasively central aortic pressure and haemodynamic indices.

These include:

1. Augmentation index adjusted to a heart rate of 75bpm (AIX@75) (unitless): this is aortic augmentation pressure expressed as a percentage of the aortic derived pulse pressure.
2. Subendocardial viability ratio/Buckberg ratio (SEVR%): area under the curve of diastolic portion divided by systolic portion of central aortic pulse wave (%).
3. Ejection duration index (%): ratio of duration of systolic ejection to total duration of cardiac cycle.

These measurements will be performed at baseline and end of study.

Statistical analysis

Each patient's change from baseline to end of study visit will be calculated. If normally distributed, the mean difference (in change from baseline) between the two groups will be calculated with 95% confidence interval. If not normally distributed, log transformation will be used if

appropriate. Alternatively, the median, interquartile range and minimum/maximum values will be reported.

If appropriate, a regression analysis will be performed including baseline and treatment group as explanatory variables, in order to estimate the means for both groups (adjusted for baseline).

6.3.8 Ankle brachial pressure index

Measurement

With the patient rested and lying supine blood pressure cuffs will be placed bilaterally on the upper arm (brachial pressure) and ankle, and inflated to 20 to 30 mmHg above systolic pressure. An ultrasound Doppler probe will be placed over the brachial, dorsalis pedis and posterior tibialis arteries and is used to detect return of the arterial signal at the highest systolic pressure.
The ABPI is calculated by dividing the ankle pressure by the brachial systolic pressure. This will be measured at baseline and end of study.

Statistical analysis

Each patient's change from baseline to end of study visit will be calculated. If normally distributed, the mean difference (in change from baseline) between the two groups will be calculated with 95% confidence interval. If not normally distributed, log transformation will be used if appropriate. Alternatively, the median, interquartile range and minimum/maximum values will be reported.
If appropriate, a regression analysis will be performed including baseline and treatment group as explanatory variables, in order to estimate the means for both groups (adjusted for baseline).

6.4 Adverse events

All reported adverse events will be tabulated by treatment group. These will be separated into two tables: unexpected adverse events and expected adverse drug reactions.

The number of participants who discontinue or who are withdrawn due to an adverse event will be tabulated. The severity and the adverse event will also be reported for these participants.

7.0 Analysis Programs

Analysis will be conducted using Stata version 11.0 or later and SPSS version 20 or later.

All analysis programmes will be stored on a networked drive (UHS or University) which can be accessed by the PI and co-investigators and RBRU/NBRC Statistician.